

SEQUENCE LISTING

(1) GENERAL INFORMATION

- (i) APPLICANT: Genencor International, Inc.
- (ii) TITLE OF THE INVENTION: ESTERASE ENZYMES, DNA ENCODING
ESTERASE ENZYMES AND VECTORS AND HOST CELLS INCORPORATING SAME
- (iii) NUMBER OF SEQUENCES: 35
- (iv) CORRESPONDENCE ADDRESS:
(A) ADDRESSEE: Genencor International, Inc.
(B) STREET: 925 Page Mill Road
(C) CITY: Palo Alto
(D) STATE: CA
(E) COUNTRY: USA
(F) ZIP: 94304-1013
- (v) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Diskette
(B) COMPUTER: IBM Compatible
(C) OPERATING SYSTEM: DOS
(D) SOFTWARE: FastSEQ for Windows Version 2.0
- (vi) CURRENT APPLICATION DATA:
(A) APPLICATION NUMBER: 08/952,445
(B) FILING DATE: 18-NOV-1997
- (vii) PRIOR APPLICATION DATA:
(A) APPLICATION NUMBER: 08/722,713
(B) FILING DATE: 30-SEP-1996
- (viii) ATTORNEY/AGENT INFORMATION:
(A) NAME: Stone, Christopher L.
(B) REGISTRATION NUMBER: 35,696
(C) REFERENCE/DOCKET NUMBER: GC362-2-US
- (ix) TELECOMMUNICATION INFORMATION:
(A) TELEPHONE: 650-846-7555
(B) TELEFAX: 650-845-6504

(2) INFORMATION FOR SEQ ID NO:1:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 32 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

Ala Ser Thr Gln Gly Ile Ser Glu Asp Leu Tyr Ser Arg Leu Val Glu
1 5 10 15

Met Ala Thr Ile Ser Gln Ala Ala Tyr Xaa Asp Leu Leu Asn Ile Pro
20 25 30

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 8 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Xaa Thr Val Gly Phe Gly Pro Tyr
1 5

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 9 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

Phe Gly Leu His Leu Xaa Gln Xaa Met
1 5

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 8 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Xaa Ile Ser Glu Asp Leu Tyr Ser
1 5

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 9 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

Tyr Ile Gly Trp Ser Phe Tyr Asn Ala
1 5

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 10 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

Gly Ile Ser Glu Asp Leu Tyr Xaa Xaa Gln
1 5 10

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 10 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

Xaa Ile Ser Glu Ser Leu Tyr Xaa Xaa Arg
1 5 10

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 7 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

Gly Ile Ser Glu Asp Leu Tyr
1 5

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 7 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

Leu Glu Pro Pro Tyr Thr Gly
1 5

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 14 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

Xaa Ala Asn Asp Gly Ile Pro Asn Leu Pro Pro Val Glu Gln
1 5 10

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 8 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

Tyr Pro Asp Tyr Ala Leu Tyr Lys
1 5

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 22 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

CGGGAATTCG CWSACCARGG AT

22

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 32 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

Variable	Mean	SD	Min	Max
Age	35.2	12.5	18	65
Gender	Male	Female	Male	Female
Marital Status	Married	Single	Married	Single
Education	High School	College	High School	College
Income	\$15,000	\$25,000	\$10,000	\$35,000
Health Status	Good	Fair	Good	Fair
Exercise Frequency	Weekly	Monthly	Weekly	Monthly
Stress Level	Low	High	Low	High
Sleep Quality	Good	Poor	Good	Poor
Dietary Habits	Healthy	Unhealthy	Healthy	Unhealthy
Work Hours	40	50	30	60
Family Size	2	3	1	4
Home Ownership	Owned	Rented	Owned	Rented
Commute Time	30	45	15	60
Job Satisfaction	High	Low	High	Low
Life Satisfaction	High	Low	High	Low
Overall Well-being	Good	Fair	Good	Fair

Val His Gly Gly Tyr Tyr Ile Gly Trp Val Ser Val Gln Asp Gln Val

1

5

10

15

(2) INFORMATION FOR SEQ ID NO:18:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 23 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

CGGGAATTCT TGGATCCRTC RTT

23

(2) INFORMATION FOR SEQ ID NO:19:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 27 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

Thr	Asp	Ala	Phe	Gln	Ala	Ser	Ser	Pro	Asp	Thr	Thr	Gln	Tyr	Phe	Arg
1				5					10					15	
Val	Thr	His	Ala	Asn	Asp	Gly	Ile	Pro	Asn	Leu					
			20					25							

(2) INFORMATION FOR SEQ ID NO:20:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 24 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

CGGGAATTCA TCCRTCRTTG CRTG

24

(2) INFORMATION FOR SEQ ID NO:21:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 27 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

Thr Asp Ala Phe Gln Ala Ser Ser Pro Asp Thr Thr Gln Tyr Phe Arg

1	5	10	15
Val	Thr	His	Ala
	Asn	Asp	Gly
	Ile	Pro	Asn
	Leu		
20	25		

(2) INFORMATION FOR SEQ ID NO:22:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 27 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

CGGGAATTCG CYTGRAAGCR TCGTCAT

27

(2) INFORMATION FOR SEQ ID NO:23:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 28 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

Met	Thr	Asp	Ala	Phe	Gln	Ala	Ser	Ser	Pro	Asp	Thr	Thr	Gln	Tyr	Phe
1				5					10					15	
Arg	Val	Thr	His	Ala	Asn	Asp	Gly	Ile	Pro	Asn	Leu				
			20				25								

(2) INFORMATION FOR SEQ ID NO:24:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 28 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

Met	Thr	Asp	Ala	Phe	Gln	Ala	Ser	Ser	Pro	Asp	Thr	Thr	Gln	Tyr	Phe
1				5					10					15	
Arg	Val	Thr	His	Ala	Asn	Asp	Gly	Ile	Pro	Asn	Leu				
			20				25								

(2) INFORMATION FOR SEQ ID NO:25:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 650 base pairs
- (B) TYPE: nucleic acid

(2) INFORMATION FOR SEQ ID NO:27:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2436 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

CCATGGTGTT GTCGATATCG GCAGTAGTCT TTGCCGAAAC GTTGAGGGTT ACAGTGATCT 60
GCGTCGGACA TACTTCGGGG AATCTACGGC GGAATATCAA AGTCTTCGGA ATATCCATAT 120
TGGGAAAGGA CAGAAGCTCC GGGGTAGTTT GATAGATGAG CTCCGGTGTA TTAAATCGGG 180
AGCTGACAGG AGTGAGCGTC ATGTAGACCA TCTAGTAATG TCAGTCGCGC GCAATTTTCGC 240
ACATGAAACA AGTTGATTTC GGGACCCCAT TGTTACATCT CTCGGCTACA GCTCGAGATG 300
TGCCTGCCGA GTATACTTAG AAGCCATGCC AGCGTGTGTG TATACGACCA AAAGTCAGGG 360
AATATGAAAC GATCGTCGGA TATTTCTTGT TTTTATCCTA AATTAGTCTT CCAGTGTTT 420
ATTTAAGAGA TAGATCCCTT CACAAACACT CATCCAACGG ACTTCTCATA CCACTCATTG 480
ACATAATTTT AAACAGCTCC AGGCGCATT AGTTCAACAT GAAGCAATTC TCCGCCAAAC 540
ACGTCCTCGC AGTTGTGGTG ACTGCAGGGC ACGCCTTAGC AGCCTCTACG CAAGGCATCT 600
CCGAAGACCT CTACAGCCGT TTAGTCGAAA TGGCCACTAT CTCCCAAGCT GCCTACGCCG 660
ACCTGTGCAA CATTCCGTCG ACTATTATCA AGGGAGAGAA AATTTACAAT TCTCAAACCTG 720
ACATTAACGG ATGGATCCTC CGCGACGACA GCAGCAAAGA AATAATCACC GTCTTCCGTG 780
GCACTGGTAG TGATACGAAT CTACAACTCG ATACTAACTA CACCCTCACG CCTTTCGACA 840
CCCTACCACA ATGCAACGGT TGTGAAGTAC ACGGTGGATA TTATATTGGA TGGGTCTCCG 900
TCCAGGACCA AGTCGAGTCG CTTGTCAAAC AGCAGGTTAG CCAGTATCCG GACTATGCGC 960
TGACTGTGAC GGGCCACAGG TATGCCCTCG TGATTTCTTT CAATTAAGTG TATAATACTC 1020
ACTAACTCTA CGATAGTCTC GGAGCGTCCC TGGCAGCACT CACTGCCGCC CAGCTGTCTG 1080
CGACATACGA CAACATCCGC CTGTACACCT TCGGCGAACC GCGCAGCGGC AATCAGGCCT 1140
TCGCGTCGTA CATGAACGAT GCCTTCCAAG CCTCGAGCCC AGATACGACG CAGTATTTCC 1200
GGGTCACTCA TGCCAACGAC GGCATCCCAA ACCTGCCCCC GGTGGAGCAG GGGTACGCC 1260
ATGGCGGTGT AGAGTACTGG AGCGTTGATC CTTACAGCGC CCAGAACACA TTTGTCTGCA 1320
CTGGGGATGA AGTGCAGTGC TGTGAGGCC AGGGCGGACA GGGTGTGAAT AATGCGCACA 1380
CGACTTATTT TGGGATGACG AGCGGAGCCT GTACATGGTG ATCAGTCATT TCAGCCTCCC 1440
CGAGTGATAC AGGAAAGATG GATGTCTCTG AGAGGGCATG CATGTACGTA TACCCGAAGC 1500
ACACTTTTTT GGTAAATCAG GACATGTAAT AAGTTCCTTC CATGAATAGA TATGGTTACC 1560
CTCACCATAA GCCTTGAGGT TGCCTTTCTC TTTTGATTGT GAATATATAT TTAAAGTAGA 1620
TGACAGATAT CTCTAAACAC CTTATCCGCT TAAACCCATC ATAGATTGTG TCACGTGATA 1680
GACCCCTTGA ATGATGAGCG AAATGTATCA GTCCCGTTTA AATCAAACCC TTTCAGCCTA 1740
GCACAGTCAG AATACACCAA CCCATTCTA AGGTAGTACT AAATATGAAT ACAGCCTAAA 1800
TGCATCGCTA TATGATCCCA TAAAGAAGCA ACAACCTTTC AGATCTCGTT TTGCGCTGCG 1860
AAGAGCTAGC TCTACCATGG TCTCAATTAT GAGTGGAGCG TTTAGTCTCG TTTAAGCCTA 1920
GCTATCTTAT AAGGACAACA CATGTACATG GGCTTACTTG TAGAGAGGTA GGATCCCGGG 1980
CTTCTTCACA TCTCGAGGAG TTGTCTACAC GTCGCGTCCA TGTCATAAGC CGGTACTCGA 2040
CGTTGTGCTG ACCGTGACCC AGACCCCTGT TGATAGCGTT GAGAAGGCC TATATTTGAA 2100
TTTCCAATCT CAGCTTTACG AAGATATGCC CATGGTGGAG GGTTAGTAAA CCGATGATGA 2160
TCGTGTGCAG CATGAGATGA GACCGTGGCC AATCCTGTTT AAATGCCAAG ACCCGCCTCC 2220
TACCACATGT AAGGCATCCG TCGGCCGCAC GTTGAATTGT GCAAATGCCG AGATCATAAA 2280
AGCGGCCACA CTTCCACGTC GGTACTGGAT GGGTTGCGCG TGGCCATACT GTGTTTTCCA 2340
TTGCGTGGGT CGTTCGTGTT ACTGCGACGC AGATTCTGTA GGCAAGGCGC AGGGCTCTCT 2400
TCTGAGGTAG AAAACACCCC ATATTAATCT GAATTC 2436

(2) INFORMATION FOR SEQ ID NO:28:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 281 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:

Met Lys Gln Phe Ser Ala Lys His Val Leu Ala Val Val Val Thr Ala
 1 5 10 15
 Gly His Ala Leu Ala Ala Ser Thr Gln Gly Ile Ser Glu Asp Leu Tyr
 20 25 30
 Ser Arg Leu Val Glu Met Ala Thr Ile Ser Gln Ala Ala Tyr Ala Asp
 35 40 45
 Leu Cys Asn Ile Pro Ser Thr Ile Ile Lys Gly Glu Lys Ile Tyr Asn
 50 55 60
 Ser Gln Thr Asp Ile Asn Gly Trp Ile Leu Arg Asp Asp Ser Ser Lys
 65 70 75 80
 Glu Ile Ile Thr Val Phe Arg Gly Thr Gly Ser Asp Thr Asn Leu Gln
 85 90 95
 Leu Asp Thr Asn Tyr Thr Leu Thr Pro Phe Asp Thr Leu Pro Gln Cys
 100 105 110
 Asn Gly Cys Glu Val His Gly Gly Tyr Tyr Ile Gly Trp Val Ser Val
 115 120 125
 Gln Asp Gln Val Glu Ser Leu Val Lys Gln Gln Val Ser Gln Tyr Pro
 130 135 140
 Asp Tyr Ala Leu Thr Val Thr Gly His Ser Leu Gly Ala Ser Leu Ala
 145 150 155 160
 Ala Leu Thr Ala Ala Gln Leu Ser Ala Thr Tyr Asp Asn Ile Arg Leu
 165 170 175
 Tyr Thr Phe Gly Glu Pro Arg Ser Gly Asn Gln Ala Phe Ala Ser Tyr
 180 185 190
 Met Asn Asp Ala Phe Gln Ala Ser Ser Pro Asp Thr Thr Gln Tyr Phe
 195 200 205
 Arg Val Thr His Ala Asn Asp Gly Ile Pro Asn Leu Pro Pro Val Glu
 210 215 220
 Gln Gly Tyr Ala His Gly Gly Val Glu Tyr Trp Ser Val Asp Pro Tyr
 225 230 235 240
 Ser Ala Gln Asn Thr Phe Val Cys Thr Gly Asp Glu Val Gln Cys Cys
 245 250 255
 Glu Ala Gln Gly Gly Gln Gly Val Asn Asn Ala His Thr Thr Tyr Phe
 260 265 270
 Gly Met Thr Ser Gly Ala Cys Thr Trp
 275 280

(2) INFORMATION FOR SEQ ID NO:29:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 2436 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:

CCATGGTGGT GTCGATATCG GCAGTAGTCT TTGCCGAAAC GTTGAGGGTT ACAGTGATCT 60
 GCGTCGGACA TACTTCGGGG AATCTACGGC GGAATATCAA AGTCTTCGGA ATATCCATAT 120
 TGGGAAAGGA CAGAAGCTCC GGGGTAGTTT GATAGATGAG CTCCGGTGTA TTAAATCGGG 180
 AGCTGACAGG AGTGAGCGTC ATGTAGACCA TCTAGTAATG TCAGTCGCGC GCAATTTTCGC 240
 ACATGAAACA AGTTGATTTT GGGACCCCAT TGTTCATCT CTGCGCTACA GCTCGAGATG 300
 TGCCTGCCGA GTATACCTAG AAGCCATGCC AGCGTGTTGT TATACGACCA AAAGTCAGGG 360
 AATATGAAAC GATCGTCGGA TATTTCTTGT TTTTATCCTA AATTAGTCTT CCAGTGGGTTT 420
 ATTTAAGAGA TAGATCCCTT CACAAACACT CATCCAACGG ACTTCTCATA CCACTCATTG 480
 ACATAATTTT AAACAGCTCC AGGCGCATT AGTTCAACAT GAAGCAATTC TCCGCCAAAC 540
 ACGTCCTCGC AGTTGTGGTG ACTGCAGGGC ACGCCTTAGC AGCCTCTACG CAAGGCATCT 600
 CCGAAGACCT CTACAGCCGT TTAGTCGAAA TGGCCACTAT CTCCCAAGCT GCCTACGCCG 660
 ACCTGTGCAA CATTCGTCG ACTATTATCA AGGGAGAGAA AATTTACAAT TCTCAAACCTG 720
 ACATTAACGG ATGGATCCTC CGCGACGACA GCAGCAAAGA AATAATCACC GTCTTCCGTG 780
 GCACTGGTAG TGATACGAAT CTACAACTCG ATACTAACTA CACCCTCACG CCTTTTCGACA 840
 CCTTACCACA ATGCAACGGT TGTGAAGTAC ACGGTGGATA TTATATTGGA TGGGTCTCCG 900
 TCCAGGACCA AGTCGAGTCG CTTGTCAAAC AGCAGGTTAG CCAGTATCCG GACTATGCGC 960
 TGACTGTGAC GGGCCACAGG TATGCCCTCG TGATTTCTTT CAATTAAGTG TATAATACTC 1020
 ACTAACTCTA CGATAGTCTC GGAGCGTCCC TGGCAGCACT CACTGCCGCC CAGCTGTCTG 1080
 CGACATACGA CAACATCCGC CTGTACACCT TCGGCGAACC GCGCAGCGGC AATCAGGCCT 1140
 TCGCGTCGTA CATGAACGAT GCCTTCCAAG CCTCGAGCCC AGATACGACG CAGTATTTCC 1200
 GGGTCACTCA TGCCAACGAC GGCATCCCAA ACCTGCCCCC GGTGGAGCAG GGGTACGCC 1260
 ATGGCGGTGT AGAGTACTGG AGCGTTGATC CTTACAGCGC CCAGAACACA TTTGTCTGCA 1320
 CTGGGGATGA AGTGCAGTGC TGTGAGGCCC AGGGCGGACA GGGTGTGAAT AATGCGCACA 1380
 CGACTTATTT TGGGATGACG AGCGGAGCCT GTACATGGTG ATCAGTCATT TCAGCCTCCC 1440
 CGAGTGTACC AGGAAAGATG GATGTCTTGG AGAGGGCATG CATGTACGTA TACCCGAAGC 1500
 ACACTTTTTC GGTAAATCAG GACATGTAAT AAGTTCCTTC CATGAATAGA TATGGTTACC 1560
 CTCACCATAA GCCTTGAGGT TGCCTTTCTC TTTTGATTGT GAATATATAT TTAAAGTAGA 1620
 TGACAGATAT CTCTAAACAC CTTATCCGCT TAAACCCATC ATAGATTGTG TCACGTGATA 1680
 GACCCCTTGA ATGATGAGCG AAATGTATCA GTCCCGTTTA AATCAAACCC TTTCAGCCTA 1740
 GCACAGTCAG AATACACCAA CCCATTCTA AGGTAGTACT AAATATGAAT ACAGCCTAAA 1800
 TGCATCGCTA TATGATCCCA TAAAGAAGCA ACAACCTTTC AGATCTCGTT TTGCGCTGCG 1860
 AAGAGCTAGC TCTACCATGG TCTCAATTAT GAGTGGAGCG TTTAGTCTCG TTTAAGCCTA 1920
 GCTATCTTAT AAGGACAACA CATGTACATG GGCTTACTTG TAGAGAGGTA GGATCCCGGG 1980
 CTTCTTCACA TCTCGAGGAG TTGTCTACAC GTCGCGTCCA TGTCATAAGC CGGTACTCGA 2040
 CGTTGTGCTG ACCGTGACCC AGACCCCTGT TGATAGCGTT GAGAAGGCCC TATATTTGAA 2100
 TTTCCAATCT CAGCTTTACG AAGATATGCC CATGGTGGAG GGTTAGTAAA CCGATGATGA 2160
 TCGTGTGCAG CATGAGATGA GACCGTGGCC AATCCTGTTC AAATGCCAAG ACCCGCCTCC 2220
 TACCACATGT AAGGCATCCG TCGGCCGCAC GTTGAATTGT GCAAATGCCG AGATCATAAA 2280
 AGCGGCCACA CTTCCACGTC GGTACTGGAT GGGTTGCGCG TGGCCATACT GTGTTTCCA 2340
 TTGCGTGGGT CGTTCGTGTT ACTGCGACGC AGATTCTGTA GGCAAGGCGC AGGGCTCTCT 2400
 TCTGAGGTAG AAAACACCCC ATATTAATCT GAATTC 2436

(2) INFORMATION FOR SEQ ID NO:30:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 32 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:

GGCCTGCAGC CCCGAAACT ACGGGTACGT CC

(2) INFORMATION FOR SEQ ID NO:31:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 35 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:31:

CGCGCTGCAG GCTCTTTCTG GTAATACTAT GCTGG

35

(2) INFORMATION FOR SEQ ID NO:32:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 36 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:32:

GGCTTAATTA ACGTGCTGGT CTCGGATCTT TGGCGG

36

(2) INFORMATION FOR SEQ ID NO:33:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 40 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:33:

GGGGCGCGCC AGATCTAGTA CCGATGTTGA GGATGAAGCT

40

(2) INFORMATION FOR SEQ ID NO:34:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 39 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:34:

GCCCAGATCT CCGCAATGAA GCAATTCTCC GCCAAACAC

39

(2) INFORMATION FOR SEQ ID NO:35:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 25 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:35:

AATAGTCGAC GGAATGTTGC ACAGG
-- 27 --
GC362-2-PCT

25

GC362-2-PCT